UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 10/043,998 | 01/11/2002 | Vitaliy S. Fain | FS-101 | 8102 |

27769 7590 03/03/2008

AKC PATENTS
215 GROVE ST.
NEWTON, MA 02466

| EXAMINER | |
|---|---|
| SMITS, TALIVALDIS IVARS | |
| **ART UNIT** | **PAPER NUMBER** |
| 2626 | |

| MAIL DATE | DELIVERY MODE |
|---|---|
| 03/03/2008 | PAPER |

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

| Office Action Summary | Application No. 10/043,998 | Applicant(s) FAIN ET AL. | |
|---|---|---|---|
| | Examiner Talivaldis Ivars Smits | Art Unit 2626 | |

*-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --*

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

1)☒ Responsive to communication(s) filed on *08 February 2008*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

4)☒ Claim(s) *1-5,9-13 and 17-25* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-5,9-13 and 17-25* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

**Application Papers**

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☒ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

## DETAILED ACTION

### *Withdrawal of Finality*

1. In response to the Advisory Action mailed 1/16/2008 and the Interview of 2/8/2008, applicant has submitted Remarks/Arguments and Interview Summary, filed 2/8/2008 arguing to overcome the art rejections, and thus for the allowability of the pending claims 1-5, 9-13, and 17-25. This has led to the withdrawal of the finality of the Office Action of 10/15/2007 and the issuance of the following corrected Final Rejection.

### *Response to Arguments*

2. Applicant's arguments with respect to the rejection(s) of independent claim(s) 1, 9, 17, 18, 19, 20, and 23 have been fully considered and are persuasive to indicate that Ju does not teach speech recognition of the recited "sound segments corresponding to words or phrases having the same spellings and different meanings", such words being called "homonyms" in examiner's cited Merriam-Webster Online Dictionary, but defined differently in Ju (per Remarks, p. 15).

However, upon further consideration, the examiner notes that these alleged "homonyms" in the Final Office Action were only one of three possible recited "at least one of" alternatives in said claims. Another alternative recited was "sound segments corresponding to words or phrases having different spellings and different meanings". The "similar sounding speech having different meanings" of Ju include words that are "pronounced alike but have different spellings" (col. 1, lines 57-58 and 38-39), referred

to as "homophones" by the examiner in said Office Action. These "homophones", whatever term is appropriate therefor, are a subset of "words or phrases having different spellings and different meanings", and thus read on this alternative.

Yes, applicant argues that Ju is referring only to the pronunciation of characters or syllables, and that "The lowest speech unit to which the notion 'meaning' can be applied is a word. (Not syllables!)" (Remarks, p. 16). However, Ju's discussion in the paragraph alluded to concerns "N-gram" language models, which are clearly stated to pertain to "n-word dependency" – thus referring to words and phrases, not to isolated characters or syllables (col. 1, lines 30-32). So, the lack of "meaning" argument fails.

Thus the Ju-based rejections in the independent claims are retained, with appropriate editing of the previous Office Action, to remove the unnecessary additional reference to words "having the same spelling and different meanings".

3. As for the argument that "There is no motivation or reason to combine" Junqua and Ju for recognizing at least one of the aforementioned spoken alternatives, because "The Junqua patent refers to processing a spoken request to control an automobile device" while "The Ju patent refers to creating a language model by associating a character string to each word" (Remarks, p. 16), the examiner disagrees.

While Ju indeed illustrates the language model by using character and syllable recognition, no such restriction is implied. In fact, Ju teaches that both "top-down" (sentence, then phrase, then word recognition) and "bottom-up" (word, then phrase, then sentence) language processing for speech recognition, "can benefit from a

language model" (col. 1, lines 19-29) such as Ju discloses "for minimizing ambiguity" (col. 1, line 57).

As for Junqua, the intended use of the disclosed natural language dialog system does not preclude its use in other contexts. If anything, one might argue that using Junqua's elegant natural language speech processor for controlling automobile devices is not obvious, because the range of expected verbal input commands is likely severely limited.

4. Applicant states that Ramaswamy has monolectic commands that are not "organized in subject areas, sub-areas, sub-sub-areas, etc." (Remarks, p. 18). Right.

However it was Thelen, not Ramaswamy, that was used to teach the claimed "*tree-like structure*, as defined by the structure in claim 1" (Remarks, p. 18). For, as indicated in the rejection of claim 1, "Thelen *et al.* teach hierarchically arranged speech recognition models, going from a more generic context to models with a more specific context" (Action, p. 5). Ramaswamy was used to teach predicting the next program module based on frequency of occurrence values, the storage thereof in a matrix being notoriously well-known so as to be an option obvious to pursue.

5. Thus, the claim rejections in the previous Office Action have been retained, *mutatis mutandis*.

## *Claim Rejections - 35 USC § 103*

1. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

2. Claims 1-4, 9-12 and 17-19 are rejected under 35 U.S.C. 103(a) as being unpatentable over Junqua (US 6,598,018) in view of Thelen *et al.* (US 6,526,380) and further in view of Ju *et al.* (US 6,934,683).

As to claim 1, Junqua teaches:

receiving a symbolic representation of a free continuous speech natural language utterance; parsing said symbolic representation of said free continuous speech natural language utterance into parsed information (a natural language interface where the input natural language is processed by a speech recognizer and supplied to a natural language parser, col. 1, lines 55-65 and col. 2, lines 12-18);

entering said parsed information into a computer instruction generator, wherein said computer instruction generator is adapted to receive inputs from a context sensitive subject area dictionary system, a context sensitive program module subdictionary system, a context sensitive argument subdictionary system and a context value subdictionary system and wherein said context sensitive subject area dictionary system comprises data organized in a plurality of subject areas, said context sensitive program

module subdictionary system comprises data organized in a plurality of program modules for each of said subject areas, said context sensitive argument subdictionary system comprise data organized in a plurality of arguments for each of said program modules and said context sensitive value subdictionary system comprises data organized in a plurality of values for reach of said arguments (entering parsed data to create computer instructions, where the computer instructions are based on the subject area (what to control, the audio or directions), what device (navigation system, radio, CD player, GPS, tape deck, or compact disk player), what command to carry out (get directions, change cd's, change volume), and how to carry out the command (directions to a point, cd to change to, what volume to change to), where each of the devices has its own context module set to control that device, having specific rules for specific functions of that device, col. 2, lines 24-43 and col. 5, lines 3-35);

determining, by accessing said context sensitive subject area dictionary system , a subject identifier, for a subject area of said parsed information (accessing the context of each of the systems to determine the subject area of the parsed information to be carried out, col. 2, lines 24-43);

determining, by accessing said context sensitive program module subdictionary system a module identifier for a program module of said subject area based upon the determined subject area identifier and the parsed information (based on the subject found and the parsed information, determining which system to command, col. 2, lines 24-43);

determining by accessing said context sensitive argument subdictionary system, an argument identifier for an argument of said program module based upon the determined identifier and the parsed information (determining what action to carry out based on the context module of the selected system, col. 2, lines 24-43);

determining, by accessing said context sensitive value subdictionary system, a value identifier for a value of said argument based upon the determined argument identifier and the parsed information (determining how to carry out the action, to what extent, col. 2, lines 24-43); and

producing computer instructions based upon the subject area identifier such that the free continuous speech natural language utterance is processed by the computer (creating computer instructions once the natural language input is received, col. 2, lines 1-22).

Junqua does not teach using a hierarchically organized context-sensitive dictionary system. However, Thelen *et al.* teach hierarchically arranged speech recognition models, going from a more generic context to models with a more specific context (col. 8, line 54 – col. 9, line 1 with Figure 4, elements 420, 422, and 424 or elements 430, 432 and 424).

Therefore, it would have been obvious to one of ordinary skill at the time of invention to have Junqua's context-sensitive dictionary system be hierarchically-organized, so as to not to have to search the entire speech recognition vocabulary but invoke the more specific models only if the more generic model gives unsatisfactory results, as Thelen *et al.* imply (col. 9, lines 12-17).

Neither Junqua nor Thelen teach that the received continuous speech natural language utterance comprises at least one of sound segments corresponding to words or phrases having the same meaning as other words or phrases corresponding to different sound segments, respectively, sound segments corresponding to words or phrses having different spellings and different meanings, or sound segments corresponding to words or phrases having a meaning that is subject area dependent.

However, Ju *et al.* (US 6,934,683) teach natural language input including sound segments having different spellings and different meanings. For, the "similar sounding speech having different meanings" of Ju include words that are "pronounced alike but have different spellings" (col. 1, lines 57-58 and 38-39). These words are a subset of "words or phrases having different spellings and different meanings", and thus read on this alternative.

It would have been obvious to one of ordinary skill at the time of invention to add the models of Ju *et al.* to Junqua and Thelen's speech recognizer to disambiguate input words having these easily confused words, as taught by Ju et al. (Title).

As to claims 9, 17, 18 and 19, Junqua teaches:

a computer readable medium comprising a set of program instructions (a natural language interface within an automobile system, where it would be inherent that the automobile system would contain a computer with instructions, since it used to control a navigation and audio system and it also inherently have a memory since it contains the ability to be updated, col. 1, lines 5-10 and col. 2, lines 55-60 and 35-43).

a receiver receiving a symbolic representation of a free continuous speech natural language utterance; a parser parsing said symbolic representation of said free continuous speech natural language utterance into parsed information (a natural language interface where the input natural language is processed by a speech recognizer and supplied to a natural language parser, col. 1, lines 55-65 and col. 2, lines 12-18);

a context sensitive subject area system dictionary system comprising data organized in a plurality of subject areas, wherein said context sensitive a subject area dictionary system is used to determine a subject area identifier for a subject area of said pares information (determining from the context module the subject of the command to be carried out, be it audio or navigational, col. 2, lines 24-43);

a context sensitive program module subdictionary system comprising data organized in a plurality of program modules for each of said subject areas and wherein said context sensitive program module subdictionary system is used to determine a module identifier for a program module of said subject area based upon the determined subject area identifier and the parsed information (determining from the subject and the parsed information what system to carry out the command on, col. 2, lines 24-43);

a context sensitive argument subdictionary system comprising data organized in a plurality of arguments for each of said program modules and where said context sensitive argument subdictionary system is used to determine an argument identifier for an argument of said program module based upon the determined module identifier and the parsed information (arguments are stored specific to each of the system that carry

out those arguments, and based on the selected system and the parsed information a selected argument is carried out, col. 2, lines 24-43);

a context sensitive value subdictionary system comprising data organized in a plurality of values for each of said arguments and wherein said context sensitive argument subdictionary system is used to determine a value identifier for a value of said argument based upon the determined argument identifier and the parsed information (data organized specific the system within the context module for that system, including how to command that system where, to what extent to change the system (volume or where to get directions to) is based on the context module and the parsed information, col. 2, lines 23-44); and

computer instructions produced based upon the subject area identifier such that the free continuous speech natural language utterance is processed by the computer (creating computer instructions once the natural language input is received, col. 2, lines 1-22).

Junqua does not teach using a hierarchically organized context-sensitive dictionary system. However, Thelen *et al.* teach hierarchically arranged speech recognition models, going from a more generic context to models with a more specific context (col. 8, line 54 – col. 9, line 1 with Figure 4, elements 420, 422, and 424 or elements 430, 432, and 424).

Therefore, it would have been obvious to one of ordinary skill at the time of invention to have Junqua's context-sensitive dictionary system be hierarchically-organized, so as to not to have to search the entire speech recognition vocabulary but

invoke the more specific models only if the more generic model gives unsatisfactory results, as Thelen *et al.* imply (col. 9, lines 12-17).

Neither Junqua nor Thelen teach that the received continuous speech natural language utterance comprises at least one of sound segments corresponding to words or phrases having the same meaning as other words or phrases corresponding to different sound segments, respectively, sound segments corresponding to words or phrases having different spellings and different meanings, or sound segments corresponding to words or phrases having a meaning that is subject area dependent.

However, Ju *et al.* (US 6,934,683) teach natural language input including sound segments having different spellings and different meanings. For, the "similar sounding speech having different meanings" of Ju include words that are "pronounced alike but have different spellings" (col. 1, lines 57-58 and 38-39). These words are a subset of "words or phrases having different spellings and different meanings", and thus read on this alternative.

It would have been obvious to one of ordinary skill at the time of invention to add the models of Ju et al. to Junqua and Thelen's speech recognizer to disambiguate input words having these easily confused words, as taught by Ju et al. (Title).

As to claims 2 and 10, Junqua teaches said subject area comprise a plurality of sub-subject areas and the context sensitive system subject area dictionary system further comprise a context sensitive sub-subject area subdictionary for each of said sub-subjects areas (the context modules have a plurality of sub-subject areas including an

audio subject, and the sub-subjects being cd player, cassette player or the radio, col. 2, lines 24-43).

As to claims 3 and 11, Junqua teaches a value identifier further comprises querying the computer system for a missing value identifier (Fig. 3b element 106).

As to claims 4 and 12, Junqua teaches: wherein determining a subject area identifier further comprises querying a user of the computer system for a missing subject area identifier; determining a module identifier further comprises querying a user of the computer system for a missing module identifier; and determining a value identifier further comprises querying a user of the computer system for a missing value identifier (if there are any missing slots that are not filled, the user is queried to supply this information, Fig. 3b, elements 94, 101, 102, 104, 106 and 108).

3. Claims 5 and 13 are rejected under 35 U.S.C. 103(a) as being unpatentable over Junqua, Thelen *et al.* and Ju *et al.* as applied to claims 1 and 9 above, and further in view of Polcyn (6,246,989).

Junqua, Thelen *et al.* and Ju *et al.* do not teach wherein, determining a subject area identifier further comprises using a previously determined value for a missing subject area identifier, determining a module identifier further comprises using a previously determined value for a missing module identifier, nor determining a value

identifier further comprises using a previously determined value for a missing value identifier.

However, Polcyn teaches receiving a natural language command from a user, and understanding the command to carry out a particular action, by determining a subject, action to be taken and argument values. Furthermore, Polcyn teaches a system that is able to determine from previous values, command information that is not understood or is missing from the current natural language input (col. 7, lines 30-40).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to combine the methods of Junqua, Thelen *et al.*, and Ju *et al.* with the teachings of Polcyn to allow a system to be updatable to contain new reference command information, as taught by Polcyn (col. 7, lines 38-40).

4. Claims 20-25 are rejected under 35 U.S.C. 103(a) as being unpatentable over Junqua, Thelen *et al.* and Ju *et al.* as applied to claims 1 and 9 above, and further in view of Ramaswamy *et al.* (6,622,119).

As to claims 20 and 23, Junqua, Thelen *et al.* and Ju *et al.* do not teach capturing a set of successfully understood free continuous speech natural language dialogs and associated program modules used to produce computer understanding, determining a frequency of occurrence value for proceeding to a next program module from a current program module, storing the frequency of occurrence values and determining the

appropriate program module selection based on choosing program modules having non-zero frequency values.

However, Ramaswamy *et al.* implies that all this was done because they teach "prompting a user for a response based on a most probable next command from the list of predicted commands" (col. 1, lines 53-55 and 60-62), which suggests such previous training to obtain the needed transition probabilities for predicting the next command in the dialog system.

That the frequency of occurrence values are stored in a matrix is not explicitly mentioned, but that is a notoriously well-known storage method, and a person with ordinary skill has good reason to pursue the known options within his or her technical grasp.

As to claims 21-22 and 24-25, Junqua, Thelen *et al.* and Ju *et al.* do not teach capturing program module or module group use frequency of occurrence information—for each step in the dialog, for proceeding to the next program module--during free continuous speech natural language dialogs.

However, Ramaswamy *et al.* suggests gathering frequency of occurrence data for commands and transitions to the next command as part of system training. So, it would have been obvious to one of ordinary skill at the time of invention to obtain this information so as to be able to modify the module structure for more efficient use of computer resources.

That the frequency of occurrence values and grouping information are stored in a matrix is not explicitly mentioned, but that is a notoriously well-known storage method for useful data, and a person with ordinary skill has good reason to pursue the known options within his or her technical grasp.

### *Conclusion*

5. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Talivaldis Ivars Smits, Ph.D., whose telephone number is 571-272-7628. The examiner can normally be reached on 8:30 a.m. to 5:00 p.m..

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on 571-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

6. Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

TĀLIVALDIS IVARS ŠMITS
PRIMARY EXAMINER

2/29/2008